Putting AI ethics to work: are the tools fit for purpose?

Year: 2021 | Citations: 157 | Authors: Jacqui Ayling, Adriane P. Chapman

Abstract

Bias, unfairness and lack of transparency and accountability in Artificial Intelligence (AI) systems, and the potential for the misuse of predictive models for decision-making have raised concerns about the ethical impact and unintended consequences of new technologies for society across every sector where data-driven innovation is taking place. This paper reviews the landscape of suggested ethical frameworks with a focus on those which go beyond high-level statements of principles and offer practical tools for application of these principles in the production and deployment of systems. This work provides an assessment of these practical frameworks with the lens of known best practices for impact assessment and audit of technology. We review other historical uses of risk assessments and audits and create a typology that allows us to compare current AI ethics tools to Best Practices found in previous methodologies from technology, environment, privacy, finance and engineering. We analyse current AI ethics tools and their support for diverse stakeholders and components of the AI development and deployment lifecycle as well as the types of tools used to facilitate use. From this, we identify gaps in current AI ethics tools in auditing and risk assessment that should be considered going forward.